

A relatively low bandwidth access connection (e.g., supporting up to about 56 kbps modem speeds) may be used to offer relatively small continuous data files (e.g., video-on-demand files), such as a two-hour video having a relatively small size of less than about 50 Mbytes). In such a case, a 72 Gbyte size disk drive has the capacity of storing about 3600 two-hour video files at the rate of about 20kbps. However, in one embodiment for supporting 10,000 simultaneous streams, about 5 disk mirrors will be needed to satisfy the required I/O operations. Therefore, one suitable storage configuration for this scenario would be one disk drive worth of content size with several mirrors. This storage configuration may be represented as “(1*72)*n”, which translates to mean “n” number of copies of one 72GB disk drive.

In the hypothetical scenario of this example, a non-resource monitoring I/O admission control algorithm similar to Resource Model Equation (18) may be suitably employed. In this embodiment, it is not necessary to employ techniques such as load balancing or workload monitoring because mirroring has made them naturally supported. Although resource monitoring techniques (e.g., IOPS monitoring) and dynamic, measurement-based I/O admission control may be implemented as described elsewhere herein, but are also not necessary in this example.

Example 8 --- General High Bandwidth Video-on-Demand Scenario

Relatively high bandwidth connections (e.g., newer last mile technologies such as xDSL lines, cable modems, fixed wireless access networks, LAN, optical networks, etc.) may be characterized as having download bandwidths of at least about 128kbps, on up to several mbps. The higher speed connections of this type are especially suitable for MPEG movies, such as those having a size of from about 1 to about 3 Gbytes.

In the embodiment of this hypothetical example, multiple information management systems (e.g., multiple content delivery systems or content routers) of a type described elsewhere herein may be deployed (e.g., on a rack) to support an aggregated throughput of from about 1000 to about 3000 streams. In this embodiment, only a small percentage of the movie library requires multiple copies, and these may be spread across the multiple information management systems,

with each information management system only needing to store at most one copy of the movie library. However, in order to support a couple of thousand movie files, each information management system may be equipped with several disk drives. Such a storage configuration may be represented as “(k*72)*1” which translates to mean “one copy of a “k” number of 72GB disk drives”.

In the hypothetical scenario of this example, a non-resource monitoring I/O admission control algorithm similar to Resource Model Equation (18) (*e.g.*, a resource model equation that considers number of storage devices NoD and a Skew type factor to estimate workload distribution of each individual disk drive) may be suitably employed where workload is substantially balanced or near balanced. For example, if the total workload is 1000 streams and the number of disk drives is 5, then a perfectly balanced average workload would be 200 streams per disk drive.

However, where workload is substantially unbalanced, a resource-monitoring capable I/O admission control algorithm may be desirable. For example, under conditions where the Skew value is 2, then the maximal workload for each individual drive may be up to $2*200 = 400$ streams. The larger the value of Skew, the greater is the workload unbalance and the smaller the total system throughput. To illustrate an extreme case, if in the system of the current hypothetical example all requests were for the same movie, then because only one disk drive has the requested content, the total number of streams would be limited to the capacity of that single disk drive, while all other drives would be idle. Furthermore, even under scenarios where movie file allocation is planned based on client access pattern (*e.g.*, more multiple disk drive space allocated to copies of “hot” or relatively popular movies), it is still possible that movie popularity may change unexpectedly, such that there still exists the potential that movie requests may suddenly be concentrated on a previously “cold” or relatively unpopular movie and the actual access Skew may exceed that previously estimated or anticipated.

Under such substantially unbalanced workload conditions, workload-monitoring may be advantageously implemented in a storage management processing engine to monitor the run-time actual workload distribution and to adjust the I/O admission control so that it will accept

relatively less requests if system workload is more skewed, and will accept relatively more requests if system workload is less skewed. Workload monitoring may be implemented, for example, using resource-monitoring capable I/O admission control algorithms such as described herein in relation to Resource Model Equation (19).

5

Example 9 --- Internet Streaming Deployment (Lower Bandwidth/Smaller File Size)

Most multimedia objects involved in current Internet streaming applications (e.g., for applications such as music exchange, news-on-demand, embedded commercials, etc.) are relatively short duration clips designed for lower-end bandwidths (e.g., from about 20kbps to about 300 kbps). The average size of these files is relatively small and, therefore the total file set size directly under the control of a storage processing engine often may fit onto a 72GB disk drive. One suitable storage organization for this exemplary embodiment would be a configuration similar to that of Example 7, and may be represented as “(1*72)*n”, namely, n mirroring of a 72 GB disk drive.

10

15

As with the scenario of Example 7, a non-resource monitoring I/O admission control algorithm similar to Resource Model Equations (18) may be suitably employed for this hypothetical situation. Resource monitoring techniques (e.g., IOPS monitoring) and/or dynamic, measurement-based I/O admission control may be implemented as described elsewhere herein, but are not necessary.

20

Example 10 --- Internet Streaming Deployment (Higher Bandwidth/Larger File Size)

Capacity improvements in the last mile infrastructure means that the number of relatively high bandwidth-capable (e.g., greater than about 300 kbps) Internet clients will continue to grow. With this trend, an increasing number of bandwidth-intensive multimedia based services will likely be offered. To provide these services, an information management system (e.g., content delivery system, content router) may be deployed under conditions where the majority of clients have high access bandwidth, and where the majority of requested objects is large. Under the conditions of this hypothetical scenario, one suitable storage organization for this exemplary

25